

(12) NACH DEM VERTRAG ÜBER DIE INTERNATIONALE ZUSAMMENARBEIT AUF DEM GEBIET DES
PATENTWESENS (PCT) VERÖFFENTLICHTE INTERNATIONALE ANMELDUNG

BERICHTIGTE FASSUNG

Serial No. 10/579,248
Group No. 1752
Confirmation No.7812

(19) Weltorganisation für geistiges Eigentum
Internationales Büro



(43) Internationales Veröffentlichungsdatum
23. Mai 2002 (23.05.2002)

PCT

(10) Internationale Veröffentlichungsnummer
WO 02/040511 A2

(51) Internationale Patentklassifikation⁷: **C07K 14/00**

(21) Internationales Aktenzeichen: PCT/EP01/12937

(22) Internationales Anmeldedatum:
8. November 2001 (08.11.2001)

(25) Einreichungssprache: Deutsch

(26) Veröffentlichungssprache: Deutsch

(30) Angaben zur Priorität:
100 55 789.9 10. November 2000 (10.11.2000) DE

(71) Anmelder (*für alle Bestimmungsstaaten mit Ausnahme von US*): **PHARMA-ZENTRALE GMBH** [DE/DE]; Loersfeldstrasse 20, 58313 Herdecke (DE).

(72) Erfinder; und

(75) Erfinder/Anmelder (*nur für US*): **HACKER, Jörg** [DE/DE]; Edith-Stein-Strasse 6, 97218 Gerbrunn (DE). **BLUM-OEHLER, Gabriele** [DE/DE]; Goethestrasse 3, 97072 Würzburg (DE). **JUNG, Günther** [DE/DE]; Ob der Grafenhalde 5, 72076 Tübingen (DE). **HANTKE, Klaus** [DE/DE]; Beckmannweg 7/1, 72076 Tübingen (DE). **PATZER, Silke** [DE/DE]; Mörikestrasse 16, 71134 Aidlingen (DE). **MORENO, Felipe** [ES/ES]; C/ Marbella N°.16, 10B, E-28034 Madrid (ES). **BAQUERO, Fernando** [ES/ES]; C/ Ardemans N°. 62, 30A, E-28028 Madrid (ES). **BAQUERO, Rosario** [ES/ES]; C/ La Baneza N°. 43 Portal 2, 90A, E-28035 Madrid (ES). **BRAVO, Daniel** [ES/ES]; C/Madero N°. 11, 10A, E-28412 Cerceda (ES). **SONNENBORN, Ulrich** [DE/DE]; Eichenweg 6, 44799 Bochum (DE). **SCHULZE, Jürgen** [DE/DE]; Alice-Bloch-Strasse 7, 14558 Bergholz-Rehbrücke (DE).

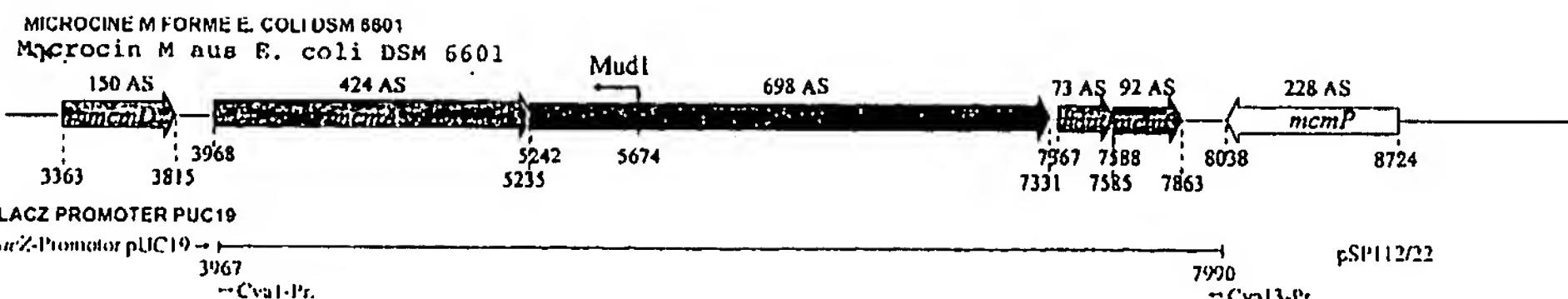
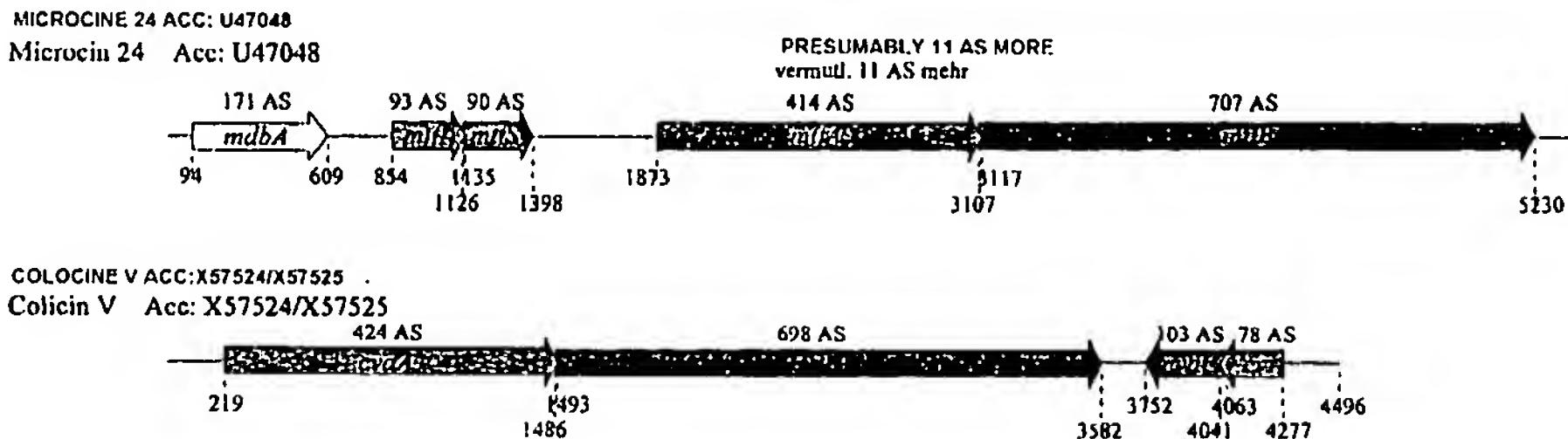
[Fortsetzung auf der nächsten Seite]

(54) Title: DNA SEQUENCE AND MICROCINE FROM ESCHERICHIA COLI STRAIN DSM 6601

(54) Bezeichnung: DNA-SEQUENZ UND MICROCIN AUS ESCHERICHIA COLI STAMM DSM 6601

ORGANISATION OF THE GENE CLUSTERS OF THE mcm REGION FROM E. COLI DSM 6601
WITH THE GENE SECTIONS NECESSARY FOR THE SYNTHESIS PROCESSING AND EXPRESSION
OF MICROCINE (BELOW) AND COMPARISON WITH THE GENE CLUSTERS FOR THE PRODUCTION
OF MICROCINE 24 (ABOVE) AND COLICINE V (MIDDLE)

Organisation des Gen-Clusters der mcm-Region aus E. coli DSM 6601 mit den
für die Microcin-Synthese, -Prozessierung und -Ausscheidung notwendigen
Genabschnitten (unten) und Vergleich mit den Gen-Clustern für die Produktion
von Microcin 24 (oben) und Colicin V (mitte)



(57) Abstract: The invention relates to a DNA sequence with the nucleotide sequence given in figure 1 and the peptide encoded thereby, derivatives of said peptide, and use of the DNA sequence, the peptide and the peptide derivative.

WO 02/040511 A2 (57) Zusammenfassung: Die Erfindung betrifft eine DNA-Sequenz mit der in Abbildung 1 dargestellten Nucleotidfolge und das dadurch codierte Peptid, Derivate des Peptids sowie die Verwendung der DNA-Sequenz, des Peptids sowie der Peptidderivate.



PROPPERT, Hans [DE/DE]; Rosenstrasse 102, 58095 Hagen (DE).

ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), OAPI-Patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

(74) **Anwalt:** SIEWERS, Gescha; Dres. Harmsen & Utescher PP., Alter Wall 55, 20457 Hamburg (DE).

Veröffentlicht:

— ohne internationalen Recherchenbericht und erneut zu veröffentlichen nach Erhalt des Berichts

(81) **Bestimmungsstaaten (national):** AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, OM, PH, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW.

(48) **Datum der Veröffentlichung dieser berichtigten Fassung:** 25. Juli 2002

(84) **Bestimmungsstaaten (regional):** ARIPO-Patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), eurasisches Patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), europäisches Patent (AT, BE, CH, CY, DE, DK,

Zur Erklärung der Zweibuchstaben-Codes und der anderen Abkürzungen wird auf die Erklärungen ("Guidance Notes on Codes and Abbreviations") am Anfang jeder regulären Ausgabe der PCT-Gazette verwiesen.

DNA-Sequenz und Microcin aus Escherichia coli Stamm DSM 6601

Die Erfindung betrifft eine DNA-Sequenz und ein neues hierdurch codiertes Microcin aus Escherichia coli Stamm DSM 6601.

5

Escherichia coli ist ein gramnegatives Bakterium, das in der menschlichen und tierischen Darmflora, aber auch extraintestinal vorkommt. E. coli tritt in zahlreichen Varianten auf, die sich hinsichtlich der Kapsel-, Oberflächen- und Flagellenantigene unterscheiden und in zahlreiche serologische Typen unterteilt werden können. Die Einordnung nach den Serotypen lässt allerdings keine Aussage über die unterschiedliche Virulenz der Keime zu. Vertreter ein- und desselben Serotyps können sowohl im menschlichen als auch im tierischen Körper ein unterschiedliches Pathogenitätspotential besitzen, das im Extremfall von avirulent bis hochgradig pathogen reichen kann. Der E. coli Stamm DSM 6601 gehört zu der Serogruppe O6:K5 und wird als nicht human und nicht tierpathogen bewertet.

20

Als Bakteriocine bezeichnet man spezifische Proteine, die von bestimmten Bakterien produziert werden und mit abtötender Wirkung gegen andere Stämme derselben oder einer nahe verwandten Bakterienart gerichtet sind. Es sind bereits über zwanzig Colicine aus E. coli bekannt, die jeweils hohe antibakterielle Spezifität aufweisen. Zu den Bakteriocinen zählen auch die sogenannten Microcine, die u.a. auch von bestimmten E. coli-Stämmen produziert werden. Bakteriocine werden, abgesehen von ihrer antibiotischen Wirkung, für die Taxonomie von Bakterienstämmen eingesetzt.

25

Es war bekannt, daß E. coli Stamm DSM 6601 antibiotische Aktivität aufweist, die dem Vorhandensein eines Microcins zugeschrieben wurde,

das im folgenden als Microcin M (McmC) bezeichnet wird. Versuche, das Peptid durch Anreicherung aus Flüssigkulturen oder durch Extraktion von Kulturen auf Agarplatten oder aus dem Agar anzureichern, brachten keine verwertbaren Ergebnisse. Es wurde daher versucht, das Microcin mit Hilfe 5 genetischer Methoden zu identifizieren. Mit dem oben genannten Stamm von *E. coli* wurde eine Mu-Mutagenese durchgeführt. Der Phage Mu kann sich nach Infektion einer Bakterienzelle wie ein Transposon verhalten und an einem beliebigen Ort in der DNA integrieren und dadurch eine Mutation hervorrufen. Im vorliegenden Fall wurde der Phage Mud 1 nach der 10 Vorschrift von Casadaban (Casadaban und Cohen, 1979, Proc. Natl. Acad. Sci. 76, 4530-4533) verwendet. Durch die Mutagenese wurde eine Mutante isoliert, die kein Microcin mehr produzierte. Aus der chromosomal DNA dieser Mutante wurde das für die Microcinsynthese kodierende DNA-Fragment isoliert und nach Insertion in den Vektor 15 pUC18 und PCR-Amplifikation kloniert und sequenziert.

Überraschenderweise wurde anhand der Sequenzanalyse festgestellt, daß die mcm-Region aus *E. coli* DSM 6601 einen dem Colicin V ähnlichen Hemmstoff kodiert, der im folgenden Microcin M (McmC) genannt wird. 20 Die Abbildung 1 zeigt die entsprechende Sequenz der mcm-Region.

Die Abbildung 1a zeigt in der Übersicht die in diesem Bereich aufgrund der DNA-Sequenz vorhergesagten Gene. Die Genprodukte McmA und McmB zeigen hohe Sequenzähnlichkeit zu den Exportproteinen für Colicin 25 V und Microcin 24. Es folgt das Gen mcml, dessen Produkt nur geringe Ähnlichkeit zu den Immunitätsproteinen der beiden obigen Microcine zeigt. Direkt im Anschluß wird das Microcin M (McmC) kodiert. Das darauf folgende Gen mcmP wird in entgegengesetzter Richtung abgelesen und wirkt aufgrund seiner Homologien zum cvpA-Gen (Colicin V Produktion)

bei der Expression von McmC mit. Stromauf vom *mcmA*-Gen liegt das Gen *mcmD*, dessen Produkt Ähnlichkeit mit den fettsäureübertragenden Proteinen (Transacylasen) aufweist, die auch in die posttranskriptionale Modifikation der *E. coli*-Hämolsine involviert sind.

5

Zur Erläuterung wird darauf hingewiesen, daß das Protein CvaB zur Familie der MDR-Proteine (Multi-Drug-Resistance) gehört, die ATP-abhängig den Export oder Import von kleinen Substanzen, aber auch den Export von Proteinen katalysieren. Es handelt sich um Proteine mit 6 oder 10 hydrophoben Helices in der Membran und einer Domäne mit einer ATP-spaltenden Aktivität, womit der Transportvorgang energetisiert wird. Die ATP-spaltende Aktivität kann auch als separates Protein vorliegen. CvaB und zehn weitere verwandte Proteine besitzen im Gegensatz zu anderen MDR-Proteinen eine N-terminale Domäne mit Cysteinproteinase-Aktivität 15 im Cytoplasma. Diese Proteinase spaltet beim Export eine kleine Leaderpeptidsequenz vom N-terminalen Ende des Colicins V ab und zwar nach der Gly-Gly Erkennungssequenz (siehe Pfeil bei dem Homologieversuch in Abbildung 1a). Eine solche Gly-Gly Sequenz findet sich auch in der abgeleiteten Sequenz für das unreife McmC, einem Peptid aus 92 Aminosäuren. Für das reife McmC-Peptid ist somit eine 20 Kettenlänge von 77 Aminosäuren vorgegeben.

Das Protein CvaA wird über eine alpha-Helix in der Cytoplasmamembran verankert und dient als Verbindungsprotein (Connector) zur äußeren Membran. In der äußeren Membran bindet CvaA an das TolC Protein. Soweit bekannt, bilden die Proteine CvaB, CvaA und TolC einen Proteinkomplex, der den Export von Colicin V über die Cytoplasmamembran und die äußere Membran in einem Schritt katalysiert. Sehr ähnliche Exportapparate hat man z.B. für das *E. coli*-

Hämolysin oder für Proteininasen aus *Erwinia*-Arten gefunden. Aufgrund der hohen Ähnlichkeit zu CvaA und CvaB dürften das McmA und McmB dieselben Funktionen beim Export und bei der Reifung von McmC erfüllen.

- 5 In den nachfolgenden Tabellen ist die prozentuale Identität der Microcine, der Exportproteine und der Immunitätsproteine sowie die prozentuale Identität im Vergleich zu anderen verwandten Proteinen zusammengestellt:

10 Tabelle 1 a: Identität der Microcine im Einzelvergleich in Prozent

	ColV	McmC	Mic24	MicE492
ColV	100	35	19	15
McmC		100	10	28
Mic24			100	45
MicE492				100

Tabelle 1 b: Identität der dem CvaA ähnlichen Exportproteine im Einzelvergleich in Prozent

15

	CvaA	McmA	MtfA
CvaA	100	97	71
McmA		100	71
MtfA			100

Tabelle 1c: Identität der dem CvaB verwandten Exportproteine in Prozent

	CvaB	McmB	MtfB
CvaB	100	89	73
McmB		100	75

MtfB		100
------	--	-----

Tabelle 1 d: Prozentuale Identität im Vergleich zu anderen McM_D verwandten Proteinen. Abkürzungen: HlyCneim ein HlyC ähnliches Protein aus *Neisseria meningitidis*; HlyCec HlyC Protein aus *Escherichia coli*

	McmD	HlyCneim	HlyCec
McmD	100	25	12
HlyCneim		100	11
HlyCec			100

Tabelle 1 e: Prozentuale Identität der Immunitätsproteine im Einzelvergleich. Abkürzungen: Cval – Immunitätsprotein von ColV, Mcml – Immunitätsprotein von Microcin McM_C, MtfI - Immunitätsprotein von Microcin M24

	Cval	Mcml	MtfI
Cval	100	23	14
Mcml		100	22
MtfI			100

Aufgrund der DNA-Analyse ergibt sich für die neue Substanz ein Peptid mit 77 Aminosäuren, das aufgrund seiner Größe den Microcinen zugerechnet wird. Dieses reife Microcin (McM_C) zeigt aber nur geringe Sequenzhomologie zu bekannten Polypeptiden wie z.B. Colicin V.

Zur Synthese und zur näheren Untersuchung des Peptids wurden überlappende Oligopeptide mit jeweils fünfzehn Aminosäuren synthetisiert, also 1 bis 15, 5 bis 20, 10 bis 25 usw., und zwar durch Oligopeptidsynthese unter Anwendung der Fmoc-Strategie auf einem Polystyrol-1% Divinylbenzolharz mit 2-Chlortriphenylmethylchlorid als Linker.

Die Gesamtsynthese des Microcin-Peptids kann in an sich bekannter Weise durch Segmentkondensation auf dem Harz unter Verwendung von drei bis vier Peptid-Segmenten und Einsatz des ABI-Peptidsynthesizers erfolgen, wobei vorzugsweise so vorgegangen wird, daß mit dem an das Harz gebundenen Segment I., bestehend aus den Aminosäuren 77 bis 54, die gereinigten Segmente II. (Aminosäuren 53 bis 24) und III. (Aminosäuren 23 bis 1) gekoppelt werden.

Die Herstellung von Lipopeptidmischungen zur Gewinnung von Antiseren erfolgt in an sich bekannter Weise, wobei Pam₃Cys-Lys-Ser-Peptide zur Immunisierung verwendet werden können.

5

Zur Beschichtung von Streptavidinplatten für ELISA-Teste werden in an sich bekannter Weise Biotinyl-Aca-Aca-Peptide verwendet, wobei diese Beschichtungen sowohl mit den entsprechenden Einzelpeptiden als auch mit einer Mischung von Peptiden durchgeführt werden können.

10

Im folgenden wird die Erfindung anhand der Beispiele näher erläutert:

Beispiel 1

15

Identifizierung und Sequenzanalyse der mcm-Region aus E. coli DSM 6601

20

Nach dem Verfahren von Casadaban et al. wurde eine Mutante, die kein Microcin mehr produzierte, im Plättentest mit dem Indikatorstamm H1941 identifiziert. Chromosomal DNA dieser Mutante wurde isoliert und mit dem Restriktionsenzym HinclI verdaut, danach mit dem mittels SmaI geschnittenen Vektor pUC18 ligiert und mit den Primern MuD1 (spezifisch für ein Mu-DNA-Ende) und UNI (spezifisch für den ligierten pUC18 Vektor) in einer PCR amplifiziert. Das DNA-Fragment wurde nach bekannten Verfahren (J. Sambrook, E.F. Fritsch, T. Maniatis; Molecular Cloning, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, 1989) kloniert und sequenziert. Es stellte sich heraus, daß dieses DNA-Fragment hohe Sequenzähnlichkeit zu dem Colicin-V-Exportprotein B aufwies. Klonierung und Sequenzanalyse des gesamten Bereiches zeigten, daß diese Region einen dem Colicin V ähnlichen Hemmstoff kodiert.

25

30

Beispiel 2

35

Synthese des Microcins

Die Aminosäuresequenz des Microcin McmC läßt sich aus der DNA der Region, die dieses Peptid kodiert, ableiten und ergibt sich aus der beigefügten Abbildung 3.

40

- Zur Synthese des Peptids werden in an sich bekannter Weise (G. Jung, Ed.; Combinatorial Peptide and Non-Peptide Libraries, Wiley VCH, Weinheim, 1996) Oligopeptide mit fünfzehn Aminosäuren synthetisiert, die überlappend die Aminosäuren 1 bis 15, 5 bis 20, 10 bis 25, 15 bis 30 usw.
- 5 aufweisen. Die Synthese erfolgt unter Verwendung von Fmoc-Schutzgruppen auf einem Polystyrol-1%-Divinylbenzolharz nach dem Merrifield-Verfahren mit 2-Chlortriphenylmethylchlorid als Linker.
- Nach hinreichend langer Synthesedauer werden in den jeweiligen
- 10 Ansätzen die Peptide vom Harz gelöst und unter Verwendung von HPLC und ES-MS-Spektroskopie untersucht. Die Sekundärstruktur kann zum Teil durch Untersuchung des Circulardichroismus festgestellt werden.
- Die Gesamtsynthese des Peptids aus siebenundsiebzig Aminosäuren
- 15 erfolgt unter Verwendung eines ABI-Peptidsynthesizers unter Kondensation von Peptidsegmenten auf dem Harz, wobei vorzugsweise mit dem harzgebundenen Peptidsegment I. (Aminosäuren 77 bis 54) angefangen wird und dann die gereinigten Segmente II.(Aminosäuren 53 bis 24) und III. (Aminosäuren 23 bis 1) gekoppelt werden.
- 20 Die freien Peptide können über HPLC gereinigt werden, wobei die Reinheit für die Derivatisierung etwa 90% betragen sollte.
- Für die Derivatisierung zur Herstellung von Lipopeptiden wird nach
- 25 publizierten Verfahren vorgegangen (K.-H. Wiesmüller, G. Jung, G. Hess; Vaccine 7: 29 – 33, 1989; G. Jung, K.-H. Wiesmüller, G. Becker, H.-J. Bühring, WG Bessler; Angewandte Chemie [Internat. Ed.] 24: 872 – 873, 1985). Die Herstellung von für ELISA-Teste geeigneten biotinylierten
- 30 Peptiden erfolgt ebenfalls nach bereits publizierten Verfahren (G. Jung, Ed.; Combinatorial Peptide and Non-Peptide Libraries, Wiley VCH, Weinheim, 1996).

Patentansprüche

1. DNA-Sequenz mit der in Abbildung 1 dargestellten Nucleotidfolge.
- 5 2. Verwendung der DNA-Sequenz nach Anspruch 1 in der mikrobiologischen Analytik und/oder Diagnostik.
- 10 3. Verwendung der DNA-Sequenz nach Anspruch 1 in der Biotechnik als Expressionsvektor oder als Bestandteil von Multi-Copy-Plasmiden zur Herstellung des antimikrobiellen Wirkstoffes.
4. Peptid mit der in Abbildung 2 dargestellten Aminosäuresequenz.
- 15 5. Verwenung des Peptids nach Anspruch 4 zur Herstellung von Antipeptid - Antikörpern.
6. Verwendung des Peptides nach Anspruch 4 als antimikrobieller Wirkstoff zu mikrobiologischen oder medizinischen Zwecken.
- 20 7. Oligopeptide mit den in Abbildung 3 dargestellten Aminosäuresequenzen.
8. Verwendung der Oligopeptide nach Anspruch 7 zur Herstellung von Antipeptid-Antikörpern.
- 25 9. Verwendung der Oligopeptide nach Anspruch 7 als antimikrobielle Wirkstoffe zu mikrobiologischen oder medizinischen Zwecken.

10. Lipopeptide nach Anspruch 4 oder 7, gekennzeichnet durch Kopplung zu Pam₃Cys-Lys-Ser-Peptiden.
 11. Peptide nach Anspruch 4 oder 7, gekennzeichnet durch Kopplung zu 5 Biotinyl-Aca-Aca-Peptiden.
 12. Verwendung der Lipopeptide nach Anspruch 10 zur Herstellung von Antiseren.
- 10 13. Verwendung der Peptide nach Anspruch 11 zur Herstellung von ELISA-Tests.

Abbildung 1 Die 51 Basenpaare lange DNA-Sequenz er kompletten mc Region aus E. coli DSM 6601

ID MCM PRELIMINARY; DNA; 5381 BP.
 SQ SEQUENCE 5381 BP; 1601 A; 971 C; 1269 G; 1540 T; 0 OTHER;
 ATAAGGGAAA TAGTAATCAT GTCTTATATA AGGGAAACCA TCAGAGGAAA AGATGAATGG
 ACTGTTATG AACAGATCGG TTTGCGGTC AGTTGTATGC TCTACAATCG TAATTACAGT
 CTGTATCCGG TGTTAACCAT TCAATACTGG ACTGAATATG CGATACAGCA TAATCAGATT
 AAATTCCGT TTGATTACAG AGGTTTCCA CTGGCGTATA TTACCTGGGC ATATCTTGAG
 GCTGATACAG AAGCACGCCCT GCTCAGGGAT CCAGAATTCA GGTTGCATCC GTCTGAATGG
 AATGAAGATG GAAGGATCTG GATTCTGGAT TTCTGTTGTA AACCAGGCTT TGGTCGAAAA
 GTTATTGACT ATCTCATACA GCTTCAGCCA TGGGGAGAAG GAGAAGTAGC ATGGTTAACG
 AGGCGAAAGA AAATTGTGAC ATACATCCCT GAGCGGCTGC ATAAAACGTA GTACCTCTGA
 AGATACAGAG ATAATTGTAA ATTACGGGGT AAATGCATCG CTGATACTAT TTTGACAGGA
 CTCTGTATTT TCTGGCTTTG ATTATAGCAT TTCCGTTAAC CAGTGAAGGT ATTACAGGAA
 AAGTGTGAGT AAGAGGAGCT ATTTTGTTC GTCAGGATGC TTTAGAAAAG AGAAAAATGA
 AGTGGCAGGG ACGGGCAATA TTACTTCCC GAAATACCACT ATGTTAACTC ATGCTGGGAA
 GCATAGTGT TATTACGGCA TTTCTGATGT TCATTATTGT TGGTACCTAT AGCCGCCGTG
 TTAATGTCAG TGGTGAGGTC ACAACCTGGC CAAGAGCTGT CAATATATAT TCAGGTGTAC
 AGGGATTGTT GTTCAGGCAG TTTGTTCATG AAGGACAGTT GATAAAAAAAA GGGGATCCTG
 TTTATCTGAT TGACATCAGT AAAAGTACAC GTAGTGGTAT TGTCACTGAT AATCATCGTC
 GGGATATAGA AAACCAGCTG GTTCGTGTGG ACAAGATTAT TTCCCGTCTG GAAGAAAGTA
 AAAAAATAAC GCTGGATAACC CTGGAAAAAC AACGCCCTGCA ATACACAGAT GCGTTTCGTC
 GCTCATCAGA TATTATACAG CGTGCAGAGG AAGGGATAAA AATAATGAAA AACAACATGG
 AGAATTACAG AAAACTATCAG ACAAAAGGGC TGATTAATAA AGATCAGTTA ACTAACCAAGG
 TGGCATTATA TTATCAGCAA CAAAACAATC TTCTTAGCCT GAGCGGACAG AACGAACAGA
 ATGCCCTGCA GATAACCACT CTGGAGAGTC AGATTTCAGAC TCAGGCAGCA GATTGGATA
 ACCGTATCTA CCAGATGGAA CTGCAACGGT ACGAGTTACA GAAAGAACTG GTTAACACTG
 ATGTGGAGGG CGAAATCATT ATCCGGCGT TGACTGACGG GAAAGTTGAC TCCCTGAGTG
 TCACTGTCGG GCAAATGGTC AATCCCAGAG ACAACCTTCT GCAGGTTATT CCTGAGAAC
 TTGAAAATTAA TTATCTTATT CTCTGGGTCC CGAATGATGC TGTTCCTTAT ATTTCGGCTG
 GTGACAAAGT GAATATTCTGT TATGAAGCCT TTCCGGCAGA AAAATTGGG CAGTTTCTG
 CTACGGTTAA AACTATATCC AGGACTCCTG CATCAACACA GGAAATGTTG ACCTATAAGG
 GAGCACCACA GAATACGCCG GGCAGCTCTG TTCCCTGGTA TAAAGTCATT GCGATGCCGT
 AAAAGCAGAT TATCAGATAT GACGAAAAAA ACCTCCCTCT GGAAAATGGA ATGAAAGCCG
 AAAGTACACT ATTTCTGGAA AAAAGGCGTA TTACCAAGTG GATGCTTCT CCTTTCTATG
 ACATGAAACA CAGTGCAACA GGACCGCTCA ATGACTAACG GGAATTTCAG ACAAATTATA
 AATCAGCTTG ATATCGCCTG GCGACGTCGT GTCCGGTTA TTCATCAGAC GGAGACCGCT
 GAATGTGGAC TGGCCTGCCT GGCAATGATA TGCAGTCATT TTGGTAAGAA TATTGACCTG
 ATATCTCTTC GTCGTAAGTT TAATCTCTCG GCCCGTGGAG CAAACCTTGC AGGAATCAAT
 GGAATAGCGG AGCAGCTGGG GATGATCACC CGGGCTCTT CACTGGAGCT GGATGAACTT
 GGTGCCCTCA AAATGCCGTG TATTCTCCAC TGGGATTTCG GTCACCTCGT CGTGCCTGGTC
 AGCGTAAAGC GTAAACCGTTA TGTACTGCAT GATCCGGCCA GAGGAAGAAG ATATCTCGGT
 CGGGAGGAAA TGAGCCGGTA TTTTACGGGC ATTGCACTTG AGGTCTGGCC CGGAGGTGAA
 TTCCGGACGG AAACCCAGCA AACCCCGATA AGTCTCCGTT CACTGATTAA CAGTATTTCAC
 GGTATTAAAA GAACACTGGC GAAAATTTC TGTCTGTCAG CTGTAATTGA AGCAATCAAT
 CTGGTAATGC CGGTGGGAAC TCAGCTGGTT ATGGATCATG CGATTCCGCC GGGGGACCGA
 GGGCTGCTGA CGCTTATTTC TGCTGGCCTG ATGTTCTTA TATTGCTCAG GGCCGCCGTG
 AGTATGCTGC GTGCATGGTC CTCACTGGTT ATGAGCACGC TCATCARAT ACAGTGGCAG
 TCGGGTCTGT TTAACCACT TCTCAGACTG CCGCTGGCAT TTTTGAACG CCGTAAATTAA
 GGTGATATCC AGTCGCGTT TGGCTCCCTT GACACTTTGA GGGCCACCTT TACCACTGT
 GTGGTTGGGG CAATCATGGA CAGTATTATG GTTGTGGGTG TTTTGTGAT GATGCTGTTA
 TATGGAGGAT ATCTTACCTG GATAGTTCTC GGTTTACCA TGCTTACCG TCTTATTCTGT
 CTGGTGACAT ACGGCTATTA CCGGCAATA TCGGAAGAAA CTCTTGTCAAG GGGGGCCCG
 GCCAGCTCCT ATTTTATGGA AAGCCTGTAT GGTATTGCCA CGGTAAAAAT CCAGGGTATG
 GTCGGGATCC GGGGAACACA CTGGCTTAAC CTGAAAATAG ATGCGATCAA TTCAGGTATT
 AAGTTAACCA GGATGGATTG GCTCTTCCGG GGAATAAATA CTTTGTGTC CGCCTGTGAT
 CAGGTGGCGA TTTTATGGCT GGGTGCAAGC CTTGTGATCG ATAATCAGAT GACAATAGGG
 ATGTTTGTGG CATTGGTTC TTTCTGTTGG CAGTTTCGG ATCGGGTTCG TTCGCTGACC
 AGTTTCTTC TTCAACTGAG AATAATGAGT CTGCATAATG AGCGCATTGC AGATATTGCA
 CTACATGAAA AGGAAGAAAA GAAACCGGAA ATTGAAATCG TTGCTGACAT GAGCCCGGTT
 TCACTGGAAA CCACTGATTT AAGCTACCGG TATGACAGCC AGTCAGCACA GGTATTCACT
 GGTCTGAATT TGTCTGTCG TCCGGGAGAA AGTGTGGCTA TAACTGGTGC CTCCGGTGC
 GGAAAAACCA CATTATGAA AGTATTATGT GGACTGTTG AACCAAGATAG TGGAAAAGTA
 CTGGTTAATG GCACGGATAT ACGTCAACTT GGAATAAATA ATTATCACCG TATGATAGCC
 TGTCTTATGC AGGACGACCG GCTATTTCG GGATCAATTG GTGAAAATAT CTGTGGGTTT
 GCAGAAGAAA CAGACGACGA ATGGATGACA GAATGTGCCA GAGCAAGTC TATTCACTGAT
 GTGATAATGA AAATGCCAAT GGGGTATGAA ACGTTAATAG GTGAACGTTT GGAAGGTCTT
 TCCGGCGGTC AAAACAGCG TATATTCACTT GCCCGAGCTT TATACCGGAA ACCTGGAATA
 TTATTTATGG ATGAGGCTAC AAGTTCTCTT GATACAGAAA GTGAACGTTT CGTTAATGCT

GCCATAAAAAA AATGAATAT CACCCGGGTG ATTATTGCAC ACAGAGAAC TACGTTGAGA
ACTGTTGACA GGATTATTTC TATTAAAAT CCACCTGGTGT AACTTGTA GGAGTTTGT
CGATGGGGGA GGTCAAGAAG GATATAAAA TAACAGTGAT TGCTTTGTT ATCAATTATC
TGTTCTTTA TATTCCGGTG TCATTATATC TTAGTTATTA CTATGGATAT AATTTTTTA
ATCTATATAT GTTTTTTTA TCACCTGTAG TTACATTTT ATCGTTGTGG TTAAACGTGA
ATTTTACTT CTTCACAAAT CTTATAGCGA AGGTGTTGAA ATGAGAAAAC TATCTGAAA
TGAAATAAAA CAAATATCTG GAGGTGACGG GAATGACGGG CAGGCAGAAT TAATTGCTAT
TGGTTCACTT GCTGGTACGT TTATTAGCCC GGGATTTGGT TCTATTGCAG GGGCTTATAT
AGGTGATAAA GTACATTCAT GGGCAACGAC TGCGACGGTT AGTCCCTCCA TGTCTCCCTC
AGGTATAGGA TTATCATCCC AGTTGGATC CGGCAGAGGT ACATCAAGTG CCTCTCGTC
TGCAGGGAGT GGAAGTTAAA CCTTATATTG TTAATGAAGC ACCTGTTACA GGTGCTTCAT
TATATGATAA TGGTGCCTAT GTAACGCTT AAGTTAGGCT CCTCCGAACG TTTATTATTT
TATGCATACC GCATAGTAAC AATGCCAGTC CGTTAATTT TTATCATATG AGATTGTAAC
AGAAAATGCT CTGTTAATGA GTATAACTGA TAGAAGCCAG ATATTGTATT CCTGCTATGA
CTGCATTCAT CGACATGTGA AGTAATATGG GCATTAACAG GCCATTGAC TTTATTTTG
CACTAATAAG AACTAACGAT ACCAGAAAGA GCATCAAAA TGTATGAAAG TCAGTATATT
GTAGATGTAA TGCAGAAAAT ATGACTGATG TTACAATAGC AGAGGTGTAA ATATTGTAT
TAAACCAGAA TTTAAAAGAG TTAAAAGAC ACCCTCGAAA TACAATCTCT TCATATACAG
GAACAAGAAG TACAGAAGAA AATATATTAA TCCATAAAAT GGCTCCGGAA AAAGACATTT
GCGAGATCAT CCATTCTTCC GTTTCCCTCA CTGCCAGCAG ATAAGGGACA AATAACTGAA
TAATCATTAT TAAAGAGAAA AGAGTAAAAA AAACATCCAG CCGAAAAGAA CCTCTTCCTA
ATCCTCTCT TTTTCGGAAG AAAATAAAAT ACAGTGGAAC AAGAATTAGA AATTCAAGCAA
GAAATAATGC TGGAACAAAGA AGTCCCCTGG ATATAAGTTC CTGCCTGTT GTAAGAAATG
CAGGAATAAA GGTAATAGAA AATGATAAAAG TAAACATTGA AAAGCAGCCG GCTGATGACT
GTTTTTTAT AGTTGCATT TTATTTTTG TCATTTCCAT T

Abbildung 1a Organisation des Gen-Clusters der mcm-Region aus E. coli DSM 6601 mit den für die Microcin-Synthese, -Prozessierung und -Ausscheidung notwendigen Genabschnitten (unten) und Vergleich mit den Gen-Clustern für die Produktion von Microcin 24 (oben) und Colicin V (mitte)

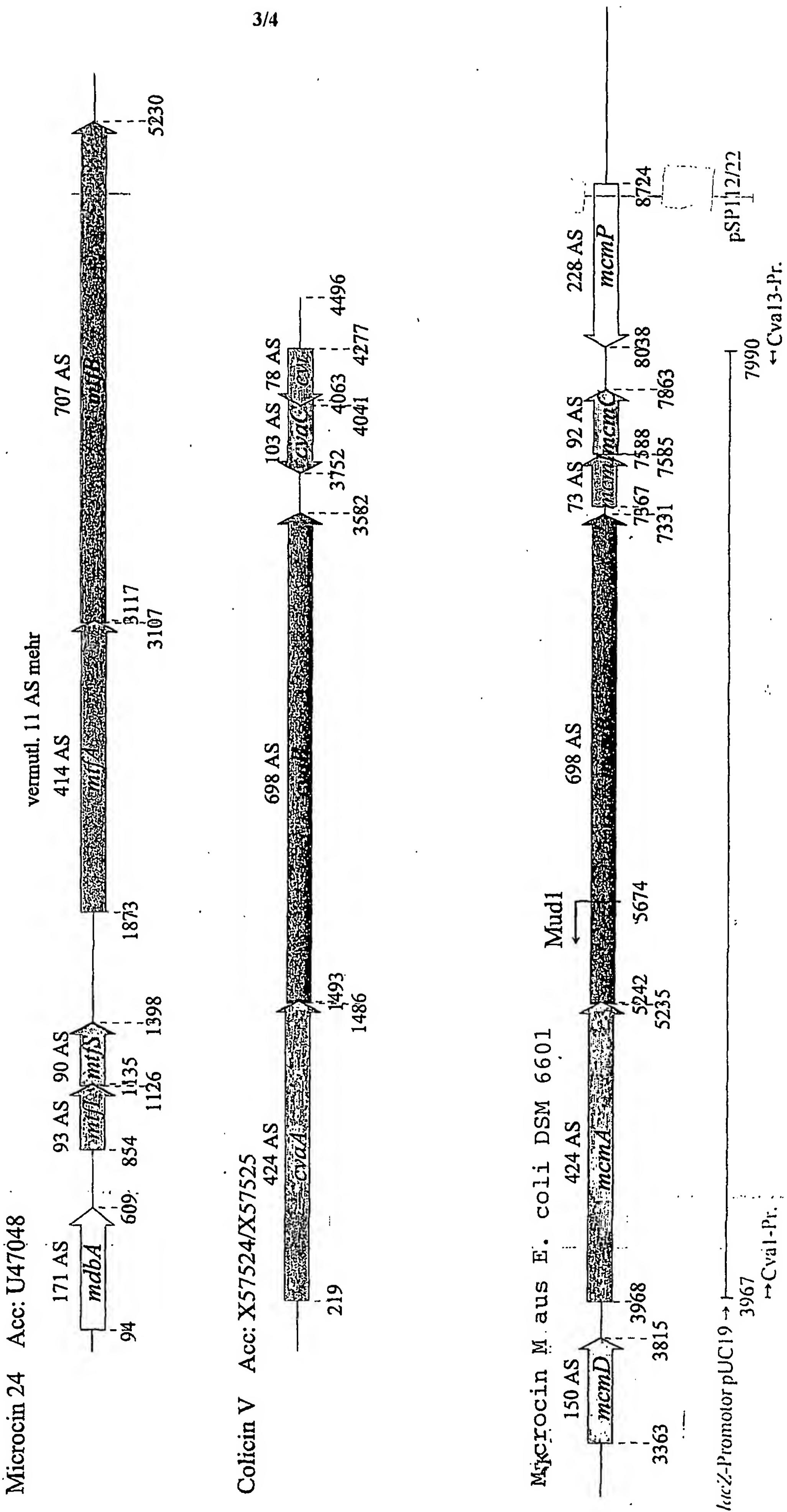


Abbildung 2 Aminosäuresequenz des antimikrobiellen Peptids (Microcin M, McmC) aus E. coli DSM 6601

Protein 1	DGNDGQAEELIAIGSLAGTFISPGFGSIAGAYIGDKVHS WATTATVSPSMSPLSGIGLSSQFGSGRGTSSASSSAGSGS
-----------	---

Abbildung 3 Überlappende Oligopeptide (Teilsequenzen) aus dem antimikrobiellen Peptid (Microcin M, McmC) aus E. coli DSM 6601

Peptide	Mol. Weight (mol/g)	Weight (mg)	mmol	Mixture	Protein
1. DGNDGQAEELIAIGSLA	1543.7	0.46	0.0003	M.1	1
2. GQAEELIAIGSLAGTFI	1560.8	0.47	0.0003	M.1	1
3. IAIGSLAITFISPGFG	1507.8	0.45	0.0003	M.1	1
4. LAGTFISPGFGSIAGA	1465.7	0.44	0.0003	M.1	1
5. ISPGFGSIAGAYIGDK	1552.8	0.47	0.0003	M.1	1
6. GSIAAGAYIGDKVHSWA	1631.3	0.49	0.0003	M.1	1
7. AYIGDKVHSWATTATV	1719.9	0.52	0.0003	M.1	1
8. KVHSWATTATVSPSMS	1639.9	0.51	0.0003	M.1	1
9. ATTATVSPSMSPLSGIG	1463.6	0.44	0.0003	M.1	1
10. VSPSMSPLSGIGLSSOF	1580.8	0.47	0.0003	M.1	1
11. SPSLIGLSSQFGSGRG	1493.6	0.45	0.0003	M.1	1
12. GLSSQFGSGRGTSSAS	1485.6	0.45	0.0003	M.1	1
13. FGSGRGTSSASSSAGS	1402.4	0.42	0.0003	M.1	1
14. SGRGTSSASSSAGSGS	1342.3	0.40	0.0003	M.1	1

SEQUENZPROTOKOLL

<110> Pharma-Zentrale GmbH

<120> DNA-Sequenz und Microcin aus Escherichia coli Stamm DSM
6601

<130> PT66/000

<140> DE 10055789.9

<141> 2000-11-10

<160> 1

<170> PatentIn Ver. 2.1

<210> 1

<211> 5381

<212> DNA

<213> Escherichia coli

<400> 1

ataaggaaaa tagtaatcat gtcttatata agggaaacca tcagaggaaa agatgaatgg 60
actgtttatg aacagatcgg ttttgcggtc agttgtatgc tctacaatcg taattacagt 120
ctgtatccgg tgttaaccat tcaatactgg actgaatatg cgatacagca taatcagatt 180
aaattcctgt ttgattcacg aggtttcca ctggcgtata ttacctggc atatcttag 240
gctgatacag aagcacgcct gctcagggat ccagaattca ggttgcattcc gtctgaatgg 300
aatgaagatg gaaggatctg gattctggat ttctgttgta aaccaggctt tggtcgaaaa 360
gttattgact atctcataca gtttcagcca tggggagaag gagaagtacg atggtaagc 420
aggcgaaaga aaattgtgac atacatccct gagcggctgc ataaaacgta gtacctctga 480
agatacagag ataattgtaa attacggggt aaatgcacg ctgatactat tttgacagga 540
ctctgtattt tctggctttg attatagcat ttccgttaac cagtgaaggt attacaggaa 600
aagtgtgagt aagaggagct attttgttc gtcaggatgc ttttagaaaac agaaaaatga 660
agtggcaggg acgggcaata ttacttcccgaataaccact atggtaatc atgctggaa 720
gcatagtgtt tattacggca tttctgtatgt tcattattgt tggtaacctat agccgcgtg 780
ttaatgtcag tggtgaggc acaacctggc caagagctgt caatataat tcaggtgtac 840
agggatttgc tgcaggcag tttgttcatg aaggacagtt gataaaaaaa gggatcctg 900
tttatctgat tgacatcagt aaaagtacac gtagtggtat tgtcaactgat aatcatcg 960
gggatataga aaaccagctg gttcgtgtgg acaagattat ttcccgctcg gaagaaagta 1020
aaaaaataac gctggatacc ctggaaaaac aacgcctgca atacacagat gcgtttcg 1080
gctcatcaga tattatacag cgtgcagagg aaggataaa aataatgaaa aacaacatgg 1140
agaattacag aaactatcag acaaaaggc tgattaataa agatcagtttta actaaccagg 1200
tggcattata ttatcagcaa caaaacaatc ttcttagcct gagcggacag aacgaacaga 1260
atgcctcgca gataaccact ctggagagtc agattcagac tcaggcagca gattttgata 1320
accgtatcta ccagatggaa ctgcaacggc acgagttaca gaaagaactg gttAACACTG 1380
atgtggaggg cgaaatcatt atccggcgt tgactgacgg gaaagttgac tccctgagt 1440
tcactgtcggt gcaaattggc aatcccggag acaaccttct gcaggttatt cctgagaaca 1500
ttgaaaattt ttatcttatt ctctgggtcc cgaatgatgc tgttccttatt atttcggctg 1560
gtgacaaagt gaatattcgt tatgaaggct ttccggcaga aaaatttggg cagtttctg 1620
ctacggtaa aactatatcc aggactcctg catcaacaca ggaaatgttg acctataagg 1680
gagcaccaca gaatacgccg ggccctctg ttccctggta taaagtcatt gcgatgcctg 1740

aaaagcagat tatcagatat gacgaaaaaa acctccctct ggaaaatgga atgaaagccg 1800
aaagtacact atttctggaa aaaaggcgta tttaccagtg gatgcttct cctttctatg 1860
acatgaaaca cagtcaaca ggaccgctca atgactaacg ggaatttcag acaaattata 1920
aatcagcttg atatgcgtcg gcgacgtcgt gttccggta ttcatcagac ggagaccgct 1980
gaatgtggac tggcctgcct ggcaatgata tgccgtcatt ttggtaagaa tattgacctg 2040
atatctttc gtcgttaagtt taatctctcg gcccggtggag caaaccttgc aggaatcaat 2100
ggaatagccg agcagctggg gatgatcacc cgggctctt cactggagct ggatgaacct 2160
ggtgcctca aaatgccgtg tattctccac tggattca gtcacttcgt cgtctggtc 2220
agcgtaaagc gtaaccgtta tgtaactgcatt gatccggcca gaggaagaag atatctcg 2280
cgggaggaaa tgagccggta tttacgggc attgcacttg aggtctggcc cggaggtgaa 2340
ttccggacgg aaaccaggca aaccgcata agtctccgtt cactgattaa cagtattac 2400
ggtattaaaa gaacactggc gaaaatttc tgtctgtcag ctgttaattga agcaatcaat 2460
ctggtaatgc cggtggaaac tcagctggtt atggatcatg cgattccggc gggggaccga 2520
gggctgctga cgcttatttc tgctggcctg atgttctta tattgctcag ggccgcggtg 2580
agtatgctgc gtgcattggc ctcactggtt atgagcacgc tcatcaatat acagtggcag 2640
tcgggtctgt ttaaccacct ttcagactg ccgtggcat ttttgaacg ccgtaaatta 2700
ggtgatatcc agtcgcgtt tggctccctt gacacttga gggccacctt taccacctgt 2760
gtgggtgggg caatcatgga cagtattatg gttgtgggtg ttttgtgat gatgctgtt 2820
tatggaggat atcttacctg gatagttctc gtttttacca tggtttacgt tcttattcgt 2880
ctggtgacat acggctatta ccggcaaata tcggaagaaa ctcttgcag gggggcccgg 2940
gccagctcct attttatgga aagcctgtat ggtattgcca cggtaaaaaat ccagggtatg 3000
gtcgggatcc ggggaacaca ctggcttaac ctgaaaatag atgcgatcaa ttcaggtatt 3060
aagttaacca ggtggatt gctttcggg ggaataaata cttttgttgc cgctgtgat 3120
caggtggcga ttttatggct gggtgcaagc cttgtgatcg ataatcagat gacaatagg 3180
atgtttgtgg catttggttc tttcgtggg cagtttcgg atcgggttgc ttcgctgacc 3240
agttttcttc ttcaactgag aataatgagt ctgcataatg aggcatttc agatattgca 3300
ctacatgaaa aggaagaaaa gaaaccggaa attgaaatcg ttgctgacat gagcccggtt 3360
tcactggaaa ccactgatt aagctaccgg tatgacagcc agtcagcaca ggtattcgt 3420
ggtctgaatt tgtctgtggc tccgggagaa agtgtggcta taactggc tccgggtgcc 3480
gaaaaaacca catatatgaa agtattatgt ggactgtttg aaccagatag tggaaaagta 3540
ctggtaatg gcacggatat acgtcaactt ggaataaata attatcaccg tatgatagcc 3600
tgtgttatgc aggacgaccg gctatttca ggatcaattc gtgaaaatat ctgtgggtt 3660
gcagaagaaa cagacgacga atggatgaca gaatgtgcca gagcaagtca tattcatgat 3720
gtgataatga aaatgccaat ggggtatgaa acgttaatag gtgaactggg ggaaggctt 3780
tccggcggtc aaaaacagcg tatattcatt gcccggactt tataccggaa acctggaaata 3840
ttatttatgg atgaggctac aagttcttt gatacagaaaa gtgaacgttt cgtaatgt 3900
gccataaaaaaa aaatgaatat cacccgggtg attattgcac acagagaaac tacgttgaga 3960
actgttgaca ggattatttc tatttaaat ccactgggt aactttgtaa ggagtttgc 4020
cgatggggga ggtcaagaag gatataaaaa taacagtgt tgctttgtt atcaattatc 4080
tgttcttttta tattccgggtg tcattatatc tttagtttata ctatggatat aattttttta 4140
atctatatat gttttttta tcacttgcgt ttacatttt atcgttgcgtt ttaaacgtga 4200
atttttactt cttcacaaat cttatagcga aggtgttgaa atgagaaaaac tatctgaaaa 4260
tgaaataaaaaa caaatatctg gaggtgacgg gaatgacggg cagggagaat taattgctat 4320
tggttcactt gctggtacgt ttatttagccc gggattttgt tctattgcag gggcttata 4380
aggtgataaaa gtacattcat gggcaacgac tgccacgggt agtccctcca tgtctccctc 4440
aggtatagga ttatcatccc agttggatc cggcagaggt acatcaagtg cctcttcgtc 4500
tgcggggagt ggaagttaaa ctttatattg ttaatgaagc acctgttaca ggtgcttcat 4560
tatatgataa tggtgccat gtaacgctt aagttaggct cctccgaacg tttattttt 4620
tatgcataacc gcatagtaac aatgccagtc cgtttaattt ttatcatatg agattgtaa 4680
agaaaatgct ctgttaatga gtataactga tagaagccag atattgtatt cctgctatga 4740
ctgcattcat cgacatgtga agtaatatgg gcattaacag gccattcgac tttattttt 4800

cactaataag aactaacat accagaaaga gcatcaaaaa tgtatgaaag tcagtatatt 4860
gtagatgtaa tgcagaaaat atgactgatg ttacaatagc agaggtgtaa atattgtcat 4920
taaaccagaa tttaaaagag ttaaaaagac accctcgaaa tacaatctt tcataatacag 4980
gaacaagaag tacagaagaa aatatattaa tccataaaaat ggctccggaa aaagacattt 5040
gcgagatcat ccattcttcc gtttcctca ctgccagcag ataagggaca aataactgaa 5100
taatcattat taaagagaaa agagtaaaaa aaacatccag ccgaaaagaa cctttccta 5160
atccttcttctt ttttcggaag aaaataaaat acagtggAAC aagaattaga aattcagcaa 5220
gaaataatgc tggaacaaga agtcccctgg atataagttc ctgcctgtt gtaagaaatg 5280
caggaataaa ggtaatagaa aatgataaaag taaacattga aaagcagccg gctgatgact 5340
gttttttat agttgcattt ttatTTTG tcatttccat t 5381